

Amendment Under 37 C.F.R. §1.116  
U.S. Patent Appln. No. 09/348,425

Docket No. 6169-125  
IBM Docket No. BOC9-1999-0036

### REMARKS

The following remarks are made in response to the final Office Action of January 22, 2003 (Office Action). In the Office Action, claims 22 through 31 have been rejected under 35 U.S.C. § 102(e) as being anticipated by U.S. Patent No. 5,799,279 to Gould *et al.* (Gould).

Prior to addressing the rejections on the art, a brief review of the Applicants' invention is appropriate. The present invention processes voice commands which include a voice command component and a dictation component wherein both components are identified from within a contiguous utterance. The voice command component is specified by a command grammar and the dictation component is freely dictated text which is embedded within the voice command structure. The voice command component is executed using at least a portion of the dictation component as an execution parameter of the voice command.

Turning to the rejections on the art, claims 22-31 have been rejected under 35 U.S.C. § 102(e) as being anticipated by Gould. In particular, with regard to claims 22 and 27, the Examiner asserts that Gould teaches the following:

- identifying a voice command having a voice command component and a dictation component within a contiguous utterance, wherein said voice command component is specified by a command grammar and said dictation component is free-form text which is not specified by said command grammar, and wherein said dictation component is embedded within said voice command; and
- executing said identified voice command component using at least a part of said dictation component as an execution parameter of said voice command.

In support, the Examiner has cited Figures 8a, 8b, 9a, 9b as well as column 5, line 13 – column 6, line 67. It is asserted that the above limitations read on "if CPU determines that user's speech is text . . . if CPU determines that user's speech was a command, then the CPU sends keystrokes or scripting language to the application to cause the application to delete the partial results from the screen and execute the command".

The cited portions of Gould, however, do not teach or suggest the features of claims 22 or 27. Claims 22 and 27 explicitly state that the voice command that is

Amendment Under 37 C.F.R. §1.116  
U.S. Patent Appln. No. 09/348,425

Docket No. 6169-125  
IBM Docket No. BOC9-1999-0036

identified includes two components, a command component and a dictation component. The portions of Gould relied on by the Examiner throughout the Office Action teach only that Gould can distinguish between dictation and commands – that is, Gould recognizes a given portion of speech as **either a command or as dictation**.

For example, the above-quoted passage (column 6, lines 18-33), describes a process wherein Gould recognizes speech. Partial results of an utterance, defined by Gould as recognized words of an utterance before the entire utterance is analyzed, can be presented on a display screen. Importantly, Gould discloses that if the CPU determines that the user speech is in fact text, the CPU is finished. The text is left upon the display screen. If, however, the user's speech is determined to be a command, the CPU recognizes that the speech is not dictation, but rather a command. The CPU then instructs the application to remove the recognized text from the display screen and instructs the application to execute the recognized text as a command. Thus, Gould teaches only that speech can be selectively recognized as **either a command or as dictated text**. **Gould does not teach that a command can be identified which includes a command component and a dictation component.**

In addition to having two separate and distinct components within the voice command, the Applicants further note that the dictation component is embedded within the voice command structure – a limitation which has been completely overlooked in the Office Action. In any case, the claim clearly indicates that each component, the command component and the dictation component, is present at the same time within the voice command.

Despite the fact that the claims call for a voice command having two different components, the Examiner continually relies upon portions of Gould which teach only a single text object, i.e. recognized speech which is determined to be either dictation or a command, for the proposition that Gould teaches "recognizing a voice command having a command component and a dictation component." For example, the Examiner states the following in the "Response to Arguments" section of the Office Action:

The examiner notes that above limitation is taught by Gould at column 6, lines 48-67; where Gould teaches, "when a user's speech is recognized

Amendment Under 37 C.F.R. §1.116  
U.S. Patent Appl. No. 09/348,425

Docket No. 6169-125  
IBM Docket No. BOC9-1999-0036

for commands and text against the same set of vocabularies, any language modeling information in the vocabularies tends to cause the system to favor the recognition of text over commands." More clearly one would understand that Gould teaches identifying a voice command component and a dictation component with a contiguous utterance in Figures 8a, 8b, 9a, and 9b[.] [A]t col. 6, lines 30-41 reads "the application being executed by the system is [a] meeting scheduler. After the system displays partial results 302 "schedule this meeting in room 507", the system determines that the utterance was a command and removes the text from the . . . display screen and executes the command by scheduling 304 the meeting in room 507. [S]imilarly, after the system displays partial results 304 "underline last three words" the system determines that that utterance was a command and removes the text from the display screen and executes the command by underlining 306 the last three words.

The above passage teaches only that speech is preliminarily recognized and presented upon a display screen. If the text is subsequently determined to be a command rather than dictation, the text is removed from the display screen and executed. That is, the text "schedule this meeting in room 507" and the text "underline last three words" are determined to be commands and not dictation. In consequence, Gould removes each phrase from the display screen and executes the commands.

Figures 8a, 8b, 9a, and 9b of the Gould specification further illustrate this aspect of Gould. In particular, Figure 8a illustrates the text "schedule this meeting in room 507" being interpreted initially and incorrectly as text. The phrase is displayed in the dictation field of the graphical user interface (GUI). Figure 8b illustrates the state of the Gould invention after the phrase "schedule this meeting in room 507" has been determined to be a command. The phrase has been removed from the dictation field and has been executed, i.e. the GUI now reflects a meeting location of "507".

Nothing in the above passage teaches or suggests that Gould recognizes a voice command which includes a dictation component and a command component, wherein the dictation component is embedded within the voice command. Rather, Gould recognizes speech as either a command or as dictation. It is noteworthy that the Examiner again refers to a single object in Gould, a text object, which is recognized exclusively as either dictation or a command, as anticipating a voice command which,

Amendment Under 37 C.F.R. §1.116  
U.S. Patent Appln. No. 09/348,425

Docket No. 6169-125  
IBM Docket No. BOC9-1999-0036

according to the literal language of the claims, includes two separate and distinct elements that coexist within the voice command.

The Examiner also cites column 6, lines 48-67 for the proposition that Gould determines a dictation component and a command component. In particular, the Examiner notes that Gould teaches that the word "bold" can be interpreted as a text component or a command component depending upon the command grammar and that the dictation component does not depend on the command grammar. The Applicants are not contending that Gould cannot recognize both voice commands and dictation. Rather, the Applicants note that Gould cannot recognize a voice command structure having a command component and a dictation component, both existing as part of the voice command structure.

The Applicants respectfully note that the literal language of the claims calls for a voice command having a command component and a dictation component, wherein the dictation component is embedded within the voice command. Hence, both components are present within the voice command. The Applicants question how the citation of a single word, the term "bold", which is interpreted in Gould as either dictation or a command, and which cannot function as both at the same time, can be inferred to anticipate two distinct elements of the Applicants' claim. Gould recognizes the term "bold" as either a command or as dictation at any given point in time. At no time, however, does Gould disclose a command comprised of both a command component and a dictation component.

Other passages cited by the Examiner within the "Response to Arguments" section of the Office Action reinforce the notion the Gould teaches distinguishing between speech that is either dictated text or a voice command, but is silent with regard to a voice command having a dictation component and a command component. For example, column 6, lines 48-67 is cited which teaches characteristics that can be used to distinguish dictated text from voice commands. For example, Gould notes how the term "bold" when used as dictation is typically followed by a noun, but when used as a command, is usually followed by a direction. The passage further notes that the system

Amendment Under 37 C.F.R. §1.116  
U.S. Patent Appln. No. 09/348,425

Docket No. 6169-125  
IBM Docket No. BOC9-1999-0036

can be weighted for favoring recognition of text as dictation rather than voice commands.

Finally, the Examiner asserts that Gould teaches "executing said identified voice command component using at least a part of said dictation component as an execution parameter of said voice command". The Examiner again cites column 6, lines 30-41 of the Gould specification. As noted above, this passage teaches only that if text that is originally regarded as dictation is subsequently determined to be a command, the text is removed from the display screen upon which the text was presented and is then executed. Again, Gould teaches only that text can be recognized as either a command or as dictation.

Claims 22 and 27 of the present invention require that the command component be executed using at least a part of the dictation component as an execution parameter of the voice command. Again, the Examiner again cites only a single phrase which is interpreted by Gould as being either exclusively dictation or exclusively a command. The Applicants fail to see how a single text object, which is either a command or dictation, anticipates a claim limitation which requires the existence of two distinct components, and further specifies that one component serves as an execution parameter.

Finally, claims 22 and 27 specify that the command component is specified by a command grammar and that **the dictation component is free-form text which is not specified by the command grammar**. Notably, both of the phrases "schedule this meeting in room 507" and "underline last three words" are commands which include no free-form text. That is, **each of these command phrases must completely conform to a fully enumerated grammar specification which is discussed at length in column 5 of the Gould specification**. In consequence, Gould does not teach a command structure as claimed in the present invention.

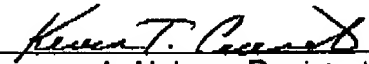
Amendment Under 37 C.F.R. §1.116  
U.S. Patent Appln. No. 09/348,425

Docket No. 6169-125  
IBM Docket No. BOC9-1999-0036

In view of the foregoing remarks, withdrawal of the 35 U.S.C. § 102(e) rejection with respect to claims 22-31 is respectfully requested. The Applicants believe that this application is now in full condition for allowance, which action is respectfully requested. The Applicants request that the Examiner call the undersigned if clarification is needed on any matter within this Amendment, or if the Examiner believes a telephone interview would expedite the prosecution of the subject application to completion.

Respectfully submitted,

Date: 4/22/03

  
Gregory A. Nelson, Registration No. 30,577  
Kevin T. Cuenot, Registration No. 46,283  
AKERMANN SENTERFITT  
222 Lakeview Avenue, Suite 400  
Post Office Box 3188  
West Palm Beach, FL 33402-3188  
Telephone: (561) 659-5000